Review Article

Basic concepts for sample size calculation: Critical step for any clinical trials!

ABSTRACT

Quality of clinical trials has improved steadily over last two decades, but certain areas in trial methodology still require special attention like in sample size calculation. The sample size is one of the basic steps in planning any clinical trial and any negligence in its calculation may lead to rejection of true findings and false results may get approval. Although statisticians play a major role in sample size estimation basic knowledge regarding sample size calculation is very sparse among most of the anesthesiologists related to research including under trainee doctors. In this review, we will discuss how important sample size calculation is for research studies and the effects of underestimation or overestimation of sample size on project's results. We have highlighted the basic concepts regarding various parameters needed to calculate the sample size along with examples.

Key words: Clinical trial; power of study; sample size

Introduction

Proper study design that is an integral component of any randomized clinical trial (i.e., the highest level of evidence available for evaluating new therapies), appears infrequently in the anesthesia literature.[1-3] Lack of sample size calculations in prospective studies and power analysis in studies with negative results have sufficiently supported this finding.[4,5] Two possible reasons for its absence in anesthesia research are followings:

1. Rare training on subjects like sample size calculation and power analysis among resident doctors.
2. Complex appearance of mathematical expression which are used for this.

To plan a research project and to understand the contents of a research paper, we need to be familiar with the fundamental concepts of medical statistics. While designing a study, we need to interact with a statistician. Understanding the basic concepts will help the anesthesiologist to interact with him in a more meaningful way. One of the pivotal aspects of planning a clinical study is the calculation of the sample size. Hence in this article, we will discuss the importance of sample size estimation for a clinical trial and different parameters that impact sample size along with basic rules for these parameters.

As we know, it is naturally neither practical nor feasible to study the whole population in any study. Hence, the sample is a set of participants (lesser in number) which adequately represents the population from which it is drawn so that true inferences about the population can be made from the results obtained. The sample size is, simply put, the number

How to cite this article: Gupta KK, Attri JP, Singh A, Kaur H, Kaur G. Basic concepts for sample size calculation: Critical step for any clinical trials!. Saudi J Anaesth 2016;10:328-31.

Gupta KK, Attri JP¹, Singh A, Kaur H, Kaur G

Department of Anesthesiology and Intensive Care, GGS Medical College and Hospital, Faridkot, ¹Department of Anesthesiology and Intensive Care, Government Medical College, Amritsar, Punjab, India

Address for correspondence: Dr. Kewal Krishan Gupta, House No. 204, Medical Campus, Faridkot - 151 203, Punjab, India.
E-mail: doc_krishan31@yahoo.co.in

of patients or experimental units in a sample. Every individual in the chosen population should have an equal chance to be included in the sample.

Why Sample Size Calculations?

The sample size is one of the first practical steps and statistical principal in designing a clinical trial to answer the research question.[6] With smaller sample size in a study, it may not be able to detect the precise difference between study groups, making the study unethical. Moreover, the results of the study cannot be generalized to the population, as this sample will be inadequate to represent the target population. On the other hand, by taking larger sample size in the study, we put more population to the risk of the intervention and also making the study unethical. It also results in wastage of precious resources and the researchers' time.

Hence, sample size is an important factor in approval/ rejection of clinical trial results irrespective of how clinically effective/ineffective, the intervention may be.

Components of sample size calculation

One of the most common questions any statistician gets asked is "How large a sample size do I need?" Researchers are often surprised to find out that the answer depends on a number of factors, and they have to give the statistician some information before they can get an answer! Thus calculating the sample size for a trial requires four basic components that are following.[7,8]

P value (or alpha)

Everybody is familiar with the term of *P* value. This is also known as level of significance and in every clinical trial we set an acceptable limit for *P* value. For example, we assume $P < 0.05$ is significant, it means that we are accepting that probability of difference in studying target due to chance is 5% or there are 5% chances of detection in difference when actually there was no difference exist (false positive results). This type of error in clinical research is also known as Type I error or alpha. Type I error is inversely proportional to sample size.

Power

Sometimes we may commit another type of error where we may fail to detect the difference when actually there is the difference. This is known as Type II error that detects false negative results, exactly opposite to mentioned above where we find false positive results when actually there was no difference. To accept or reject null hypothesis by adequate power, acceptable limit for the false negative rate must be decided before conducting the study.

In another term, Type II error is the probability of failing to find the difference between two study groups when actually a difference exist and it is termed as beta (β). By convention, maximum acceptable value for β in bio-statistical literature is 0.20 or a 20% chance that null hypothesis is falsely accepted. The "power" of the study then is equal to (1-β) and for a β of 0.2, the power is 0.8, which is the minimum power required to accept the null hypothesis. Usually, most of clinical trial uses the power of 80% which means that we are accepting that one in five times (i.e., 20%) we will miss a real difference. The power of a study increases as the chances of committing a Type II error decrease.

The effect size

Effect size (ES) is the minimal difference that investigator wants to detect between study groups and is also termed as the minimal clinical relevant difference. We can estimate the ES by three techniques that is, pilot studies, previously reported data or educated guess based on clinical experiences.

To understand the concept of ES, here we take one example. Suppose that treatment with Drug A results in a reduction of mean blood pressure (MBP) by 10 mm of Hg and with Drug B results in a reduction of MBP by 20 mm of Hg. Then, absolute ES will be 10 mm of Hg in this case. ES can be expressed as the absolute or relative difference. As in above example, relative difference/reduction with drug intervention is 10/20 or 50%. In other word, for continuous outcome variables the ES will be numerical difference and for binary outcome e.g., effect of drug on development of stress response (yes/no), researcher should estimate a relevant difference between the event rates in both treatment groups and could choose, for instance, a difference of 10% between both the groups as ES. In statistics, the difference between the value of the variable in the control group and that in the test drug group is known as ES. Even a small change in the expected difference with treatment has a major effect on the estimated sample size, as the sample size is inversely proportional to the square of the difference. For larger ES, smaller sample size would be needed to prove the effect but for smaller ES, sample size should be large.

The variability

Finally for the sample size calculation, researcher needs to anticipate the population variance of a given outcome variable which is estimated by means of the standard deviation (SD). Investigators often use an estimate obtained from information in previous studies because the variance is usually an unknown quantity. For a homogenous population, we need smaller sample size as variance or SD will be less in this population. Suppose for studying

the effect of diet program A on the weight, we include a population with weights ranging from 40 to 105 kg. Now, it is easy to understand that the SD in this group will be more and we would need a larger sample size to detect a difference between interventions, else the difference between the study groups would be concealed by the inherent difference between them because of the SD. If, on the other hand, we take a sample from a population with weights between 60 and 80 kg we would naturally get a more homogenous group, thus reducing the SD and, therefore, the sample size.

Other factors affecting the sample size calculation are dropout rate and underlying event rate in the population.

Dropout rate

Dropout rate means that estimate of a number of subjects those can leave out the study/clinical trial due to some reason. Normally the sample size calculation will give a number of study subjects required for achieving target statistical significance for a given hypothesis. However in clinical practice, we may need to enroll more subjects to compensate for these potential dropouts.[9]

If n is the sample size required as per formula and if d is the dropout rate then adjusted sample size N1 is obtained as N1 = $n/(1-d)$.

Event rate in population

Prevalence rate or underlying event rate of the condition under study in population is very important while calculating sample size. It is usually estimated from previous literature. For example studying the association of smoking and brain tumor, the prevalence rate for a brain tumor in studying population should be known prior to the study. Sometimes, we have to readjust the sample size after starting the trial because of unexpectedly low event rate in the population.

Examples of sample size calculations
Example 1: Comparing two means
A placebo-controlled randomized trial proposes to assess the effectiveness of Drug A in preventing the stress response to laryngoscopy. A previous study showed that there is the average rise of 20 mm of Hg in systolic blood pressure during laryngoscopy (with SD of 15 mm of Hg). What should be the sample size to study the difference in mean systolic blood pressure between two groups with a significant $P < 0.05$ and power of study of 90%.

Level of significance = 5%, Power = 80%, $Z\alpha$ = Z is constant set by convention according to accepted α error and Z

$(1-\beta)$ = Z is constant set by convention according to power of study which is calculated from Table 1.

Formula of calculating sample size is:

$n = 2 (Z\alpha + Z [1-\beta])^2 \times SD^2/d^2$
$Z\alpha = 1.96$, $Z (1-\beta) = 1.28$, SD = 15, d (effect size) = 20
So $n = 2 (1.96 + 1.28)^2 \times 15^2/20^2 = 11.82$

Twelve individuals in each group should be recruited in the study.

Example 2: Comparing two proportions
A study is planned to check the effectiveness of dexmedtomidine in preventing postoperative shivering. From previous studies, it was found that incidence of postoperative shivering is 60%. Reduction in the incidence of shivering up to 20% will be considered significant for the effectiveness of the drug. What should be the sample size for the study with alpha value, that is, Type I error of 5% (0.05) and power of 95%.

The parameters for sample size calculation are:

P1 = 60%, P2 = 20%, Effect size (d) = 40%, $Z\alpha$ = 1.96, Z $(1-\beta)$ = 1.64
Average of P1 and P2 will be p = P1 + P2/2 60 + 20/2 = 40
q = 1-p = 1-40 = 60%
Formula will be: $n = 2 (Z\alpha + Z [1-\beta])^2 \times p \times q/d^2$
$n = 2 (1.96 + 1.64)^2 \times 40 \times 60/(40)^2 = 38.88$

In this example, we need 40 persons to be included in the each group for the study.

Some basic rules for on sample size estimations are
1. Power — It should be ≥80%. Sample size increases as power increases. Higher the power, lower the chance of missing a real effect.[10]
2. Level of significance — It is typically taken as 5%. The sample size increases as level of significance decreases.[10]

Table 1: Showing constant values for convention values of α and β values

Z values for conventional values of alpha (α)	
Alpha	$Z\alpha$
0.05	1.96
0.01	2.58
Z values for conventional values of beta (β)	
Beta	Z $(1-\beta)$
0.20	0.842
0.10	1.28
0.05	1.64
0.01	2.33

3. Clinically meaningful difference — To detect a smaller difference, one needs a sample of large sample size and vice a versa.[10]
4. *Post-hoc* power calculations — After the study is conducted, one should not perform any "*post-hoc*" power calculations. The power should always be calculated prior to a study to determine the required sample size, since it is the prestudy probability that the study will detect a minimum effect regarded as clinically significant. Once the effect of the study is known, investigators should use the 95% confidence interval (CI) to express the amount of uncertainty around the effect estimate instead of power.[11,12]

Reporting of sample size

According to the CONSORT statement, sample size calculations should be reported and justified in all published RCTs. Many studies only include statements like "we calculated that the sample size in each treatment group should be 150 at an alpha of 0.05 and a power of 0.80." However, such a statement is almost meaningless because it neglects the estimates for the effect of interest and the variability. The best way to express sample size from IDEAL clinical trial should be "A clinically significant effect of 10% or more over the 3 years would be of interest. Assuming 3-year survival rates in the control group and the intervention group of 64% and 74% respectively, with a two-sided significance of 0.05 and a power of 0.8, a total of 800-1000 patients will be required."[13]

Conclusion

The sample size is first and important step in planning a clinical trial and any negligence in its estimation may lead to rejection of an efficacious drug and an ineffective drug may get approval. Although techniques for sample size calculation are described in various statistical books, performing these calculations can be complicated and it is desirable to consult an experienced statistician in estimation of this vital study parameter. To have meaningful dialogue with the statistician every research worker should be familiar with the basic concepts of sample size calculations.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References

1. Pua HL, Lerman J, Crawford MW, Wright JG. An evaluation of the quality of clinical trials in anesthesia. Anesthesiology 2001;95:1068-73.
2. Moodie PF, Craig DB. Experimental design and statistical analysis. Can Anaesth Soc J 1986;33:63-5.
3. Villeneuve E, Mathieu A, Goldsmith CH. Power and sample size calculations in clinical trials from anesthesia journals. Anesth Analg 1992;74:S337.
4. Mathieu A, Villeneuve E, Goldsmith CH. Critical appraisal of methodological reporting in the anesthesia literature. Anesth Analg 1992;74:S195.
5. Fisher DM. Statistics in anesthesia. In: Miller RD, editor. Anesthesia. 4th ed. New York: Churchill Livingstone; 1994. p. 782-5.
6. Kirby A, Gebski V, Keech AC. Determining the sample size in a clinical trial. Med J Aust 2002;177:256-7.
7. Kadam P, Bhalerao S. Sample size calculation. Int J Ayurveda Res 2010;1:55-7.
8. Dhulkhed VK, Dhorigol MG, Mane R, Gogate V, Dhulkhed P. Basics statistical concepts for sample size estimation. Indian J Anaesth 2008;52:788-93.
9. Sakpal TV. Sample size estimation in clinical trial. Perspect Clin Res 2010;1:67-9.
10. Brasher PM, Brant RF. Sample size calculations in randomized trials: common pitfalls. Can J Anaesth 2007;54:103-6.
11. Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. Ann Intern Med 1994;121:200-6.
12. Greenland S. On sample-size and power calculations for studies using confidence intervals. Am J Epidemiol 1988;128:231-7.
13. Cooper BA, Branley P, Bulfone L, Collins JF, Craig JC, Dempster J, *et al.* The initiating dialysis early and late (IDEAL) study: Study rationale and design. Perit Dial Int 2004;24:176-81.